X

RECEIVED

MAR 2 9 2002

TECH CENTER 1600/2900

1600

#21
4-2-02
P.7.

RAW SEQUENCE LISTING   DATE: 03/26/2002
PATENT APPLICATION: US/09/357,675C   TIME: 14:13:26

Input Set : **A:\PTO.AMC.txt**
Output Set: **N:\CRF3\03262002\I357675C.raw**

P.6

```
 4 <110> APPLICANT: Croce, Carlo M.
 6 <120> TITLE OF INVENTION: Nitrilase Homologs
 9 <130> FILE REFERENCE: CR001.NP001
11 <140> CURRENT APPLICATION NUMBER: 09/357,675C
12 <141> CURRENT FILING DATE: 1999-07-20
14 <150> PRIOR APPLICATION NUMBER: 60/093,350
15 <151> PRIOR FILING DATE: 1998-07-20
17 <160> NUMBER OF SEQ ID NOS: 31
19 <170> SOFTWARE: FastSEQ for Windows Version 4.0
21 <210> SEQ ID NO: 1
22 <211> LENGTH: 1416
23 <212> TYPE: DNA
24 <213> ORGANISM: Homo sapiens
26 <220> FEATURE:
27 <221> NAME/KEY: misc_feature.
28 <222> LOCATION: (19)...(19)
29 <223> OTHER INFORMATION: n=a
31 <400> SEQUENCE: 1
```

RECEIVED ENTERED

RECEIVED

MAR 2 9 2002

TECH CENTER 1600/2900

```
32 gcccactcgc tgcggcctnt ctggctccag accgccctcc ggatcggacc ctgcgaatgg    60
33 ttttggctat atcttcatgt aggacctact ccctatcccg tcggccgcgg ctgggcttca   120
34 tcaccaggcc tcctcacaga ttcctgtccc ttctgtgtcc tggactccgg atacctcaac   180
35 tctcagtact ttgtgctcag cccaggccca gagccatggc tatctcctct tcctcctgcg   240
36 aactgcccct ggtggctgtg tgccaggtaa catcgacgcc agacaagcaa cagaacttta   300
37 aaacatgtgc tgagctggtt cgagaggctg ccagactggg tgcctgcctg gctttcctgc   360
38 ctgaggcatt tgacttcatt gcacgggacc ctgcagagac gctacacctg tctgaaccac   420
39 tgggtgggaa acttttggaa gaatacaccc agcttgccag ggaatgtgga ctctggctgt   480
40 ccttgggtgg tttccatgag cgtggccaag actgggagca gactcagaaa atctacaatt   540
41 gtcacgtgct gctgaacagc aaaggggcag tagtggccac ttacaggaag acacatctgt   600
42 gtgacgtaga gattccaggg caggggccta tgtgtgaaag caactctacc atgcctgggc   660
43 ccagtcttga gtcacctgtc agcacaccag caggcaagat tggtctagct gtctgctatg   720
44 acatgcggtt ccctgaactc tctctggcat tggctcaagc tggagcagag atacttacct   780
45 atccttcagc ttttggatcc attacaggcc cagcccactg ggaggtgttg ctgcgggccc   840
46 gtgctatcga aacccagtgc tatgtagtgg cagcagcaca gtgtggacgc caccatgaga   900
47 agagagcaag ttatggccac agcatggtgg tagacccctg gggaacagtg gtggcccgct   960
48 gctctgaggg gccaggcctc tgccttgccc gaatagacct caactatctg cgacagttgc  1020
49 gccgacacct gcctgtgttc cagcaccgca ggcctgacct ctatggcaat ctgggtcacc  1080
50 cactgtctta agacttgact tctgtgagtt tagacctgcc cctcccaccc ccaccctgcc  1140
51 actatgagct agtgctcatg tgacttggag gcaggatcca ggcacagctc ccctcacttg  1200
52 gagaaccttg actctcttga tggaacacag atgggctgct gggaaagaa  actttcacct  1260
53 gagcttcacc tgaggtcaga ctgcagtttc agaaaggtgg aattttatat agtcattgtt  1320
54 tatttcatgg aaactgaagt tctgctgagg gctgagcagc actggcattg aaaaatataa  1380
55 taatcataaa gtcaaaaaaa aaaaaaaaaa aaaaaa                            1416
57 <210> SEQ ID NO: 2
```

RAW SEQUENCE LISTING   DATE: 03/26/2002
PATENT APPLICATION: US/09/357,675C   TIME: 14:13:26

Input Set : **A:\PTO.AMC.txt**
Output Set: **N:\CRF3\03262002\I357675C.raw**

```
 58 <211> LENGTH: 23
 59 <212> TYPE: DNA
 60 <213> ORGANISM: Homo sapiens
 62 <400> SEQUENCE: 2
 63 tctgaaactg cagtctgacc tca                                    23
 65 <210> SEQ ID NO: 3
 66 <211> LENGTH: 21
 67 <212> TYPE: DNA
 68 <213> ORGANISM: Homo sapiens
 70 <400> SEQUENCE: 3
 71 caggcacagc tcccctcact t                                      21
 73 <210> SEQ ID NO: 4
 74 <211> LENGTH: 20
 75 <212> TYPE: DNA
 76 <213> ORGANISM: Homo sapiens
 78 <220> FEATURE:
 79 <221> NAME/KEY: misc_feature
 80 <222> LOCATION: (0)...(0)
 81 <223> OTHER INFORMATION: n=a, g, c or t
 83 <400> SEQUENCE: 4
 84 gtngtnccng gncaygtngt                                        20
 86 <210> SEQ ID NO: 5
 87 <211> LENGTH: 26
 88 <212> TYPE: DNA
 89 <213> ORGANISM: Homo sapiens
 91 <220> FEATURE:
 92 <221> NAME/KEY: misc_feature
 93 <222> LOCATION: (0)...(0)
 94 <223> OTHER INFORMATION: n=a,c,g, or t and y=c or t
 96 <400> SEQUENCE: 5
 97 acrtgnacrt gyttnacngt ytgngc                                 26
 99 <210> SEQ ID NO: 6
100 <211> LENGTH: 21
101 <212> TYPE: DNA
102 <213> ORGANISM: Drosophila melanogaster
104 <400> SEQUENCE: 6
105 gcgcctttgt ggcctcgact g                                      21
107 <210> SEQ ID NO: 7
108 <211> LENGTH: 21
109 <212> TYPE: DNA
110 <213> ORGANISM: Drosophila melanogaster
112 <400> SEQUENCE: 7
113 cggtggcgga agttgtctgg t                                      21
115 <210> SEQ ID NO: 8
116 <211> LENGTH: 20
117 <212> TYPE: DNA
118 <213> ORGANISM: Caenorhabditis elegans
120 <400> SEQUENCE: 8
121 gtggcggctg ctcaaactgg                                        20
```

RAW SEQUENCE LISTING   DATE: 03/26/2002
PATENT APPLICATION: US/09/357,675C   TIME: 14:13:26

Input Set : **A:\PTO.AMC.txt**
Output Set: **N:\CRF3\03262002\I357675C.raw**

```
123 <210> SEQ ID NO: 9
124 <211> LENGTH: 21
125 <212> TYPE: DNA
126 <213> ORGANISM: Caenorhabditis elegans
128 <400> SEQUENCE: 9
129 tcgcgacgat gaacaagtcg g                                    21
131 <210> SEQ ID NO: 10
132 <211> LENGTH: 19
133 <212> TYPE: DNA
134 <213> ORGANISM: Homo sapiens
136 <400> SEQUENCE: 10
137 gccctccgga tcggaccct                                      19
139 <210> SEQ ID NO: 11
140 <211> LENGTH: 20
141 <212> TYPE: DNA
142 <213> ORGANISM: Homo sapiens
144 <400> SEQUENCE: 11
145 gacctactcc ctatcccgtc                                     20
147 <210> SEQ ID NO: 12
148 <211> LENGTH: 21
149 <212> TYPE: DNA
150 <213> ORGANISM: Homo sapiens
152 <400> SEQUENCE: 12
153 gctgcgaagt gcacagctaa g                                   21
155 <210> SEQ ID NO: 13
156 <211> LENGTH: 24
157 <212> TYPE: DNA
158 <213> ORGANISM: Homo sapiens
160 <400> SEQUENCE: 13
161 aaactgaagc ctctttcctc tgac                                24
163 <210> SEQ ID NO: 14
164 <211> LENGTH: 20
165 <212> TYPE: DNA
166 <213> ORGANISM: Homo sapiens
168 <400> SEQUENCE: 14
169 tgggcttcat caccaggcct                                     20
171 <210> SEQ ID NO: 15
172 <211> LENGTH: 22
173 <212> TYPE: DNA
174 <213> ORGANISM: Homo sapiens
176 <400> SEQUENCE: 15
177 ctgggctgag cacaaagtac tg                                  22
179 <210> SEQ ID NO: 16
180 <211> LENGTH: 21
181 <212> TYPE: DNA
182 <213> ORGANISM: Homo sapiens
184 <400> SEQUENCE: 16
185 gcttgtctgg cgtcgatgtt a                                   21
187 <210> SEQ ID NO: 17
```

```
188 <211> LENGTH: 36
189 <212> TYPE: DNA
190 <213> ORGANISM: Homo sapiens
192 <400> SEQUENCE: 17
193 tgacgtcgac atatgtcaac tctagttaat accacg                          36
195 <210> SEQ ID NO: 18
196 <211> LENGTH: 25
197 <212> TYPE: DNA
198 <213> ORGANISM: Homo sapiens
200 <400> SEQUENCE: 18
201 tgggtacctc gactagctta tgtcc                                      25
203 <210> SEQ ID NO: 19
204 <211> LENGTH: 147
205 <212> TYPE: PRT
206 <213> ORGANISM: Homo sapien
208 <220> FEATURE:
209 <223> OTHER INFORMATION: Xaa is an unknown amino acid
211 <400> SEQUENCE: 19
212 Met Ser Phe Arg Phe Gly Gln His Leu Ile Lys Pro Ser Val Val Phe
213 1               5                  10               15
214 Leu Lys Thr Glu Leu Ser Phe Ala Leu Val Asn Arg Lys Pro Val Val
215           20              25              30
216 Pro Gly His Val Leu Val Cys Pro Leu Arg Pro Val Glu Arg Phe His
217           35              40              45
218 Asp Leu Arg Pro Asp Glu Val Ala Asp Leu Phe Gln Thr Thr Gln Arg
219       50              55              60
220 Val Gly Thr Val Val Glu Lys His Phe His Gly Thr Ser Leu Thr Phe
221 65              70              75              80
222 Ser Xaa Gln Asp Gly Pro Glu Ala Gly Gln Thr Val Lys His Val His
223           85              90              95
224 Val His Val Leu Pro Arg Lys Ala Gly Asp Phe His Arg Asn Asp Ser
225           100             105             110
226 Ile Tyr Glu Glu Leu Gln Lys His Asp Lys Glu Asp Phe Pro Ala Ser
227       115             120             125
228 Trp Arg Ser Glu Glu Glu Glu Ala Ala Glu Ala Ala Ala Leu Arg Val
229       130             135             140
230 Tyr Phe Gln
231 145
234 <210> SEQ ID NO: 20
235 <211> LENGTH: 150
236 <212> TYPE: PRT
237 <213> ORGANISM: murine
239 <400> SEQUENCE: 20
240 Met Ser Phe Arg Phe Gly Gln His Leu Ile Lys Pro Ser Val Val Phe
241 1               5                  10               15
242 Leu Lys Thr Glu Leu Ser Phe Ala Leu Val Asn Arg Lys Pro Val Val
243           20              25              30
244 Pro Gly His Val Leu Val Cys Pro Leu Arg Pro Val Glu Arg Phe Arg
245           35              40              45
```

RAW SEQUENCE LISTING   DATE: 03/26/2002
PATENT APPLICATION: US/09/357,675C   TIME: 14:13:26

Input Set : **A:\PTO.AMC.txt**
Output Set: **N:\CRF3\03262002\I357675C.raw**

```
246 Asp Leu His Pro Asp Glu Val Ala Asp Leu Phe Gln Val Thr Gln Arg
247     50              55              60
248 Val Gly Thr Val Val Glu Lys His Phe Gln Gly Thr Ser Ile Thr Phe
249 65              70              75              80
250 Ser Met Gln Asp Gly Pro Glu Ala Gly Gln Thr Val Lys His Val His
251             85              90              95
252 Val His Val Leu Pro Arg Lys Ala Gly Asp Phe Pro Arg Asn Asp Asn
253             100             105             110
254 Ile Tyr Asp Glu Leu Gln Lys His Asp Arg Glu Glu Glu Asp Ser Pro
255         115             120             125
256 Ala Phe Trp Arg Ser Glu Lys Glu Met Ala Ala Glu Ala Glu Ala Leu
257     130             135             140
258 Arg Val Tyr Phe Gln Ala
259 145             150
262 <210> SEQ ID NO: 21
263 <211> LENGTH: 327
264 <212> TYPE: PRT
265 <213> ORGANISM: Homo sapien
267 <400> SEQUENCE: 21
268 Met Leu Gly Phe Ile Thr Arg Pro Pro His Arg Phe Leu Ser Leu Leu
269 1               5               10              15
270 Cys Pro Gly Leu Arg Ile Pro Gln Leu Ser Val Leu Cys Ala Gln Pro
271             20              25              30
272 Arg Pro Arg Ala Met Ala Ile Ser Ser Ser Ser Cys Glu Leu Pro Leu
273         35              40              45
274 Val Ala Val Cys Gln Val Thr Ser Thr Pro Asp Lys Gln Gln Asn Phe
275     50              55              60
276 Lys Thr Cys Ala Glu Leu Val Arg Glu Ala Ala Arg Leu Gly Ala Cys
277 65              70              75              80
278 Leu Ala Phe Leu Pro Glu Ala Phe Asp Phe Ile Ala Arg Asp Pro Ala
279             85              90              95
280 Glu Thr Leu His Leu Ser Glu Pro Leu Gly Gly Lys Leu Leu Glu Glu
281             100             105             110
282 Tyr Thr Gln Leu Ala Arg Glu Cys Gly Leu Trp Leu Ser Leu Gly Gly
283         115             120             125
284 Phe His Glu Arg Gly Gln Asp Trp Glu Gln Thr Gln Lys Ile Tyr Asn
285     130             135             140
286 Cys His Val Leu Leu Asn Ser Lys Gly Ala Val Val Ala Thr Tyr Arg
287 145             150             155             160
288 Lys Thr His Leu Cys Asp Val Glu Ile Pro Gly Gln Gly Pro Met Cys
289             165             170             175
290 Glu Ser Asn Ser Thr Met Pro Gly Pro Ser Leu Glu Ser Pro Val Ser
291             180             185             190
292 Thr Pro Ala Gly Lys Ile Gly Leu Ala Val Cys Tyr Asp Met Arg Phe
293         195             200             205
294 Pro Glu Leu Ser Leu Ala Leu Ala Gln Ala Gly Ala Glu Ile Leu Thr
295     210             215             220
296 Tyr Pro Ser Ala Phe Gly Ser Ile Thr Gly Pro Ala His Trp Glu Val
297 225             230             235             240
```

VARIABLE LOCATION SUMMARY   DATE: 03/26/2002
PATENT APPLICATION: US/09/357,675C   TIME: 14:13:27

Input Set : **A:\PTO.AMC.txt**
Output Set: **N:\CRF3\03262002\I357675C.raw**

Use of n's or Xaa's(NEW RULES):
Use of n's and/or Xaa's have been detected in the Sequence Listing.
Use of <220> to <223> is MANDATORY if n's or Xaa's are present.
in <220> to <223> section, please explain location of n or Xaa, and which
residue n or Xaa represents.

Seq#:1; N Pos. 19
Seq#:4; N Pos. 3,6,9,12,18
Seq#:5; N Pos. 6,15,18,24
Seq#:19; Xaa Pos. 82
Seq#:25; Xaa Pos. 6